



HD28

.M414

no. 3465
-92

WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

Inspection for Circuit Board Assembly

Philippe B. Chevalier
and
Lawrence M. Wein

WP# 3465-92-MSA

September, 1992

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

Inspection for Circuit Board Assembly

Philippe B. Chevalier
and
Lawrence M. Wein

WP# 3465-92-MSA

September, 1992

M.I.T. LIBRARIES
OCT 20 1992
RECEIVED

INSPECTION FOR CIRCUIT BOARD ASSEMBLY

Phillipe B. Chevalier

Operations Research Center, M.I.T.

and

Lawrence M. Wein

Sloan School of Management, M.I.T.

Abstract

Several stages of tests are performed in circuit board assembly, and each test consists of one or more noisy measurements. We consider the problem of jointly optimizing the allocation of inspection and the testing policy; that is, at which stages should a board be inspected, and at these stages, whether to accept or reject a board based on noisy test measurements. The objective is to minimize the expected costs for testing, repair, and defective items shipped to customers. We observe that the optimal testing policy is very sensitive to some input data that is extremely hard to estimate accurately in practice. Hence, we do not pursue an optimal solution, although it could be calculated with considerable effort. Rather, we derive a closed form testing policy that is Pareto optimal with respect to the expected number of type I and type II errors, and then find an optimal inspection allocation given this testing policy. We also describe an application of the model to an industrial facility that is using it, and find that the proposed policy significantly improves upon the facility's historical policy.

September 1992

INSPECTION FOR CIRCUIT BOARD ASSEMBLY

Philippe B. Chevalier

Operations Research Center, M.I.T.

and

Lawrence M. Wein

Sloan School of Management, M.I.T.

1. Introduction

This paper considers the problem of inspection in a circuit board assembly plant, which came to our attention while working with a Hewlett-Packard facility. Recent technological advances have given rise to circuit boards of increasing complexity, and consequently testing has become one of the most challenging aspects of circuit board manufacturing. Inspection costs can account for over half of the total manufacturing cost, and hence optimizing the utilization of inspection resources is a crucial task.

The assembly of circuit boards is performed in a single manufacturing stage, which is followed by several successive inspection stages. Assembled circuit boards have to be inspected for manufacturing defects and defective components. The inspection process at each stage includes three different activities: testing, diagnosis and repair. In testing, one or more measurements are taken from a board, and a decision is made whether to accept or reject a board. The defect on a rejected board is identified during diagnosis, and is corrected during repair. Diagnosis is often the most difficult and time consuming task and the degree of difficulty depends on the board type and the type of test that is performed. The main focus of this paper is on testing, and henceforth repair refers to diagnosis as well as repair.

The attractiveness of a test depends on its cost, diagnostic power, coverage and measurement errors. The cost for performing a test on a particular board depends on the type of that board. The *diagnostic power* of a test is the amount of information that a test

provides for diagnosis, and it strongly influences the time required and the cost for a repair. Defects are considered to be of different types, and the coverage of a test refers to the range of defects that can be detected, and the extent to which they can be detected. For example, in Figure 2, the defect types linked to defective assembly, such as opens and shorts, are very well covered at the first inspection stage, whereas defective components are much harder to detect and some components can only be detected at the last inspection stage. The outcome of a test is one or more measurements, and the accuracy of these measurements determines the prevalence of type I (false reject) and type II (false accept) errors.

We consider two interrelated decisions. The first decision is to decide at which stage(s) to inspect a board; this problem is known in the literature as the *inspection allocation problem*. At the stages where inspection is performed, the *testing policy* decides whether to accept or reject each board based on the noisy measurements obtained from the test. The joint decision of inspection allocation and testing will be referred to as an *inspection policy*. The optimization problem is to find an inspection policy to minimize the total expected cost, which includes costs for testing, repair and defective items shipped to customers. Since we assume that every defective board is repaired, no scrapping cost is included.

We make one crucial assumption that allows for a tractable analysis: the probability distribution of the true value of a measurement at a particular stage is independent of the inspection policy employed at previous stages. If the measurement under consideration is trying to detect a type i defect, then one would expect that the presence of previous inspections related to type i defects, as well as the use of tight acceptance intervals for these tests, would lead to a narrower distribution of the true measurement value. The exact nature of this dependency can be very complex since the same quantity is not measured at different stages. The quantities measured at later stages are typically an aggregation of quantities measured earlier, and the nature of this aggregation might either dampen or amplify the effect of a defect. Even if a model of this dependency was developed, data collection would be extremely difficult. Typically, the only available data is the noisy measurements under the

existing inspection policy; experiments would probably need to be performed under various inspection policies to gather the necessary data. In our case study, the derivation of an optimal testing policy was primarily an issue at in circuit testing (many of the other tests had binary results), which is the first test undertaken, and consequently the independence assumption was not an issue.

We encountered a formidable obstacle to implementing an optimal solution to this problem: the optimal testing policy is very sensitive to the probability distributions of the measurement errors, which are exceedingly difficult to estimate to the required degree of precision. Hence, the considerable computational effort required to derive the optimal inspection policy did not appear to be worthwhile. Instead, we found a testing policy with several important features: (i) the cutoff limits for acceptance and rejection that characterize the testing policy are expressed in closed form; consequently, the test engineers at Hewlett-Packard found the policy intuitively appealing, and easy to understand and use; (ii) the policy is Pareto optimal with respect to the expected number of type I and type II errors incurred; and (iii) the cutoff limits should be close to the optimal cutoff limits if reasonably accurate parameter estimates are used.

Given this testing policy, the optimal inspection allocation policy is then easily calculated. Since the cardinality of the action space is small and the dimensionality of the continuous state space can be large, we employed an exhaustive search procedure instead of a dynamic programming algorithm.

The paper contains a short case study of the Hewlett-Packard facility, where our model is being applied. The case study includes a description of the facility, the methodology for gathering the data, which has been disguised, and the proposed policy. Relative to this facility's historical inspection policy, our limited numerical results suggest that the proposed allocation policy can reduce total inspection costs by 10-20%, and the proposed testing policy can reduce costs by roughly 5%. The proposed inspection policy is in the process of being implemented, and we are not yet in a position to report on the cost reduction realized by

the facility.

Two by-products of our analysis are of significant practical value. First, we determine the cost reduction that would be achieved by reducing the measurement noise of the test equipment. This quantity can help circuit board manufacturers evaluate new test equipment, and can assist test equipment manufacturers focus their R&D efforts and market their equipment. Although inspection is necessary in the context of circuit board assembly, quality improvement efforts are also vital to a company's success. We also derive the marginal benefit from reducing various types of defects. These quantities can be used to focus quality improvement efforts in an economic fashion.

Many papers have been published on the optimal allocation of inspection in multistage serial systems. Early studies on this problem (see, for example, White 1966 and Lindsay and Bishop 1967) assumed perfect inspection (i.e., no type I or type II errors). Later papers allowed for imperfect inspection (Eppen and Hurst 1974, Yum and McDowell 1981 and Garcia-Diaz et al. 1984), where one determines the number of times that a test should be repeated. A survey of work published on inspection allocation can be found in Raz (1986). Recently, Villalobos et al. (1992) studied a dynamic version of the same problem, where inspection of an item at a particular stage can depend on the result of the inspection of that item at previous stages. Our paper appears to be the first to address the presence of distinct defect types, the joint optimization of inspection allocation and testing, and the application of a model to an industrial facility.

Section 2 presents a detailed formulation of the problem, which is then analyzed in Section 3. A case study is presented in Section 4, and conclusions are drawn in Section 5.

2. Problem Formulation

A typical flow chart of the assembly of circuit boards is presented in Figure 1. The main manufacturing stage is circuit board assembly, where all the components are soldered onto the printed circuit boards. At the system assembly, the different boards are plugged

into the final product. The *in circuit test* takes a measurement from each of the individual components soldered on a board, and the *functional test* assesses the response of a board to simulated working conditions. At the *system test*, each board is tested as part of a complete system. Many variants of the configuration displayed in Figure 1 are possible. For example, there could be multiple levels of each test, or the system assembly step could be performed in several stages, and a test could be performed on each subassembly; on the other hand, some tests might not be present.

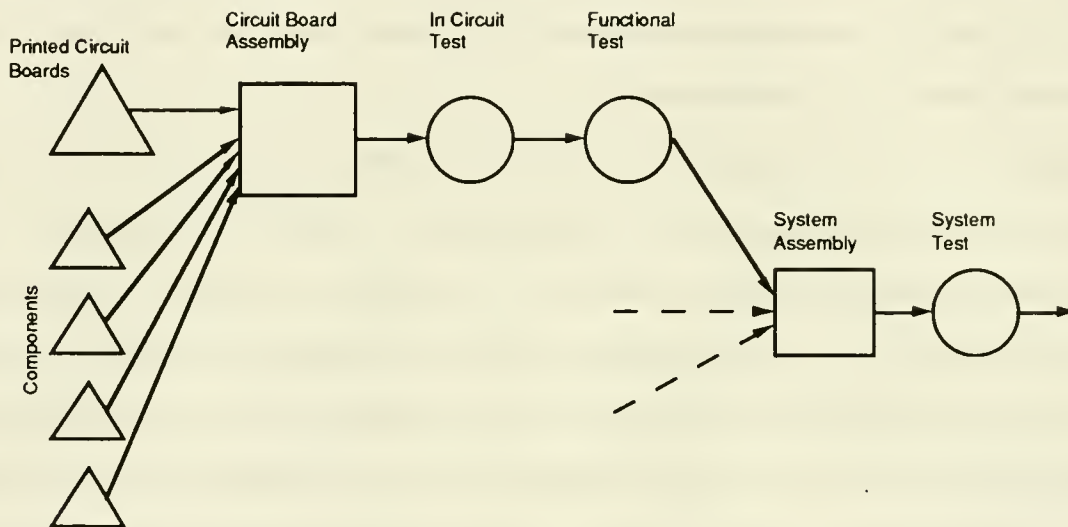


Figure 1. Flow chart of circuit board assembly.

As mentioned earlier, an important characteristic of each test is its coverage. We will assume that for each type of defect, the successive tests that are performed on the circuit boards have an increasing coverage. Figure 2 illustrates this property, which we call *hierarchical test coverage*, for the example introduced above. Since they cover more defect types, successive tests tend to be more complex and more expensive to perform. However, as the complexity of the test increases, the precision of the information obtained from the test decreases. Consequently, type I and type II errors are more prevalent, and diagnosis takes longer and has to be performed by more qualified personnel. The hierarchical assumption in test coverage holds for the circuit board assembly systems we have encountered. This assumption also holds in traditional manufacturing settings, where additional work is per-

formed between successive tests, and each test measures the cumulative functionality of the manufactured item.

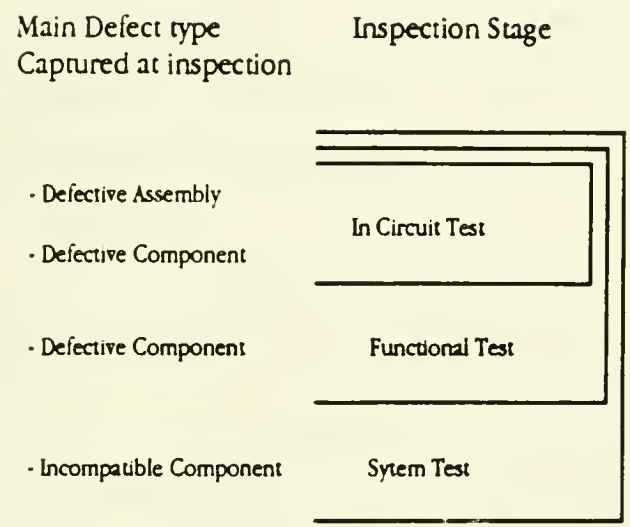


Figure 2. Hierarchical test coverage.

In this section, the problem is formulated for a single board type in isolation. In Section 3.3 we discuss how to account for the dependencies between different board types.

A test consists of a series of up to several thousand measurements. These measurements relate to different components on a board, or to different aspects of the overall functional performance of a board. Most of these measurements are subject to some noise, and this measurement error is quite sensitive to small changes in the board design or the manufacturing process. The testing policy at a stage specifies an interval for each measurement such that a board will be accepted if every measurement lies inside its interval, and rejected otherwise. The presence of measurement errors makes it impossible to completely eliminate type I and type II errors. Larger intervals will lead to the acceptance of more boards, which will reduce the number of good boards falsely rejected but increase the number of defective boards falsely accepted. Similarly, smaller intervals will have the opposite effect.

For tests that have binary results, only one testing policy needs to be considered. Examples of such situations are systemwide functional tests and tests for opens and shorts. For the purpose of generality, we allow the model to have a choice of testing policies at each

stage. However, the model easily accommodates the case where the set of possible testing policies at one or more stages is reduced to a single policy.

We assume that each measurement can detect only one type of defect. For $i = 1, \dots, I$ and $n = 1, \dots, N$, let K_{in} be the number of measurements taken at stage n that detect type i defects, and let v_{ink}^m be the value obtained from the k^{th} measurement for type i defects at stage n . Then

$$v_{ink}^m = v_{ink}^t + \epsilon_{ink},$$

where v_{ink}^t is the true measurement value and ϵ_{ink} is the measurement error. We assume that these two random variables are independent, and that the measurement errors are independent across different measurements. These independence assumptions appear to hold in practice. As mentioned in the Introduction, we also make the simplifying assumption that the distribution of v_{ink}^t is independent of the inspection policy at previous stages.

Let $p_{ink}(x)$ be the conditional probability that the true value v_{ink}^t is inside G_{ink} , which is the interval in which the true value should be for the proper functioning of the board, given the measurement value; that is,

$$p_{ink}(x) = \Pr[v_{ink}^t \in G_{ink} \mid v_{ink}^m = x].$$

The function $p_{ink}(x)$ can be expressed solely in terms of the problem inputs as

$$p_{ink}(x) = \frac{\int_{G_{ink}} e_{ink}(x - y) \xi_{ink}(y) dy}{\int_{-\infty}^{+\infty} e_{ink}(x - y) \xi_{ink}(y) dy}, \quad (1)$$

where $\xi_{ink}(x)$ is the probability density function of the true value of the quantity measured, and $e_{ink}(x)$ is the density function of the measurement error.

The testing policy for defect type i at stage n is defined by $T_{in} = \{(L_{ink}, U_{ink}), k = 1, \dots, K_{in}\}$, where the k^{th} measurement for type i defects at stage n is accepted if it lies inside $[L_{ink}, U_{ink}]$. Define

$$\alpha_{ink}(L_{ink}, U_{ink}) = \Pr[v_{ink}^t \in G_{ink} \text{ and } v_{ink}^m \notin [L_{ink}, U_{ink}]]$$

and

$$\beta_{ink}(L_{ink}, U_{ink}) = \Pr \left[v_{ink}^t \notin G_{ink} \text{ and } v_{ink}^m \in [L_{ink}, U_{ink}] \right],$$

where $\alpha_{ink}(L_{ink}, U_{ink})$ is the probability that the measurement is rejected although the true value of the component is good, and $\beta_{ink}(L_{ink}, U_{ink})$ is the probability that the measurement is accepted while the true value of the component is bad. If we let $f_{ink}(x)$ be the density function of the measured values, then

$$\alpha_{ink}(L_{ink}, U_{ink}) = \int_{-\infty}^{L_{ink}} p_{ink}(x) f_{ink}(x) dx + \int_{U_{ink}}^{\infty} p_{ink}(x) f_{ink}(x) dx \quad (2)$$

and

$$\beta_{ink}(L_{ink}, U_{ink}) = \int_{L_{ink}}^{U_{ink}} (1 - p_{ink}(x)) f_{ink}(x) dx. \quad (3)$$

Let $\alpha_{in}(T_{in})$ be the expected number of false defects of type i per board at stage n under testing policy T_{in} , and let $\beta_{in}(T_{in})$ be the expected number of defects of type i present on a board at stage n that are not detected at that stage. It follows that

$$\alpha_{in}(T_{in}) = \sum_{k=1}^{K_{in}} \left[\int_{-\infty}^{L_{ink}} p_{ink}(x) f_{ink}(x) dx + \int_{U_{ink}}^{\infty} p_{ink}(x) f_{ink}(x) dx \right] \quad (4)$$

and

$$\beta_{in}(T_{in}) = \sum_{k=1}^{K_{in}} \int_{L_{ink}}^{U_{ink}} (1 - p_{ink}(x)) f_{ink}(x) dx. \quad (5)$$

For technical reasons that will become apparent later, we assume that $p_{ink}(x)$ and $f_{ink}(x)$ are continuous unimodal functions for all i, n and k .

In addition to deciding upon a testing policy, we also need to choose the inspection allocation policy at each stage, which specifies whether or not to test the boards at that stage. Notice that additional stages can be artificially added to consider partial testing options at particular stages. The objective is to minimize the total expected cost of testing, repair, and defects leaving the plant; this total cost will often be referred to as the *total inspection cost*. We consider a per unit testing cost t_n at stage n , which typically includes operator time, test engineering, equipment depreciation and maintenance, and various overhead costs. A possible benefit, or negative cost, of testing is that it will lead to quicker learning and hence

process improvements; however, in our case study, we did not attempt to incorporate this benefit into the per unit testing cost. Also, no fixed testing cost is included in the model. The repair cost r_{in} is the total cost incurred to diagnose and repair a type i defect on a board at stage n . We assume that all defective boards are repaired, and that the same repair cost is incurred, whether the defect is a real defect or a false defect. The cost f of a defect on a board that leaves the plant includes the cost of a field repair, the cost of the analysis and repair of the defective board that comes back to the plant, and a cost measuring the customer's loss of goodwill. We assume that the cost is per defect and not per defective system, which simplifies the analysis. However, if there are two or more defects in a system, then it is unlikely that these defects would be detected by the customer at the same time. Moreover, since the number of defects leaving the plant is very low, multiple defects on the same system are very unlikely; consequently, the results obtained would be very similar if a cost was incurred per defective system. We also assume that the cost f does not depend on the type of defect on a board. Although the more general case can be easily accommodated, this assumption seems practical, since the costs of lost goodwill and visiting the repair site dominate the other costs, and are independent of the defect type.

As a consequence of the hierarchical test coverage assumption, more defects become detectable at each inspection stage. We model this as if, on average, d_{in} new defects of type i appear per board at stage n . Let δ_{in} denote the average number of defects of type i per board that leave stage n . Notice that δ_{in} depends on the inspection policy at all earlier stages, and d_{in} and δ_{in} are expected values. The distributions of the random variables from which these expected values are derived do not matter because all the costs in the model are linear and dynamic allocation policies (i.e., policies where the decision to inspect depends upon what was observed earlier on the board) are not considered. If inspection is performed at stage n and testing policy T_{in} is used, then the expected cost of testing and repair at that stage is

$$t_n + \sum_{i=1}^I r_{in} (\delta_{i,n-1} + d_{in} + \alpha_{in}(T_{in}) - \beta_{in}(T_{in})),$$

and the expected number of defects of type i per board leaving stage n is

$$\delta_{in} = \beta_{in}(T_{in}).$$

On the other hand, if no inspection is performed at stage n , then no cost is incurred at that stage. The expected number of type i defects per board leaving stage n in this case is

$$\delta_{in} = \delta_{i,n-1} + d_{in}.$$

The optimization problem is to decide whether to inspect at each stage, and if so, what testing policy T_{in} to use, if such a choice exists, to minimize the expected total inspection cost,

$$\sum_{n=1}^N \left(t_n + \sum_{i=1}^I r_{in} (\delta_{i,n-1} + d_{in} + \alpha_{in}(T_{in}) - \beta_{in}(T_{in})) \right) + f \sum_{i=1}^I \delta_{iN}. \quad (6)$$

3. Analysis

In this section, we analyze the problem formulated in Section 2. The testing problem is addressed in Section 3.1 and the inspection allocation policy is numerically derived in Section 3.2. Various extensions of the basic problem are described in the last three subsections.

3.1. The Testing Policy

This subsection contains three main results. First, we derive a set of testing policies that is optimal with respect to the tradeoff between the expected number of type I and type II errors per board in the sense of Pareto optimality (i.e., it is impossible to reduce one quantity without increasing the other). This result enables us to restrict our attention to these policies when searching for an optimal testing policy. Second, three rather weak additional assumptions are made to sharpen the first result. However, we observe that the optimal testing policy is very sensitive to certain input data that are difficult to estimate in practice. Hence, we conclude this subsection by proposing a closed form policy that is Pareto optimal and easy to implement and understand.

The minimum expected total inspection cost from stage n until exiting the plant, which is denoted by $J_n(\delta_{1n}, \delta_{2n}, \dots, \delta_{In})$, satisfies the dynamic programming optimality equations

$$J_n(\rho_1, \rho_2, \dots, \rho_I) = \text{Min} \left[J_{n+1}(\rho_1 + d_{1n}, \rho_2 + d_{2n}, \dots, \rho_I + d_{In}), \right. \\ \left. t_n + \min_{T_{in}} \left[\sum_{i=1}^I r_{in} (\rho_i + d_{in} + \alpha_{in}(T_{in}) - \beta_{in}(T_{in})) \right. \right. \\ \left. \left. + J_{n+1}(\beta_{1n}(T_{in}), \beta_{2n}(T_{in}), \dots, \beta_{In}(T_{in})) \right] \right] \text{ for } n = 1, \dots, N, \quad (7)$$

and

$$J_{N+1}(\rho_1, \rho_2, \dots, \rho_I) = f \sum_{i=1}^I \rho_i, \quad (8)$$

where $(\rho_1, \rho_2, \dots, \rho_I)$ are dummy variables.

The second minimization in (7) can be rewritten more concisely as

$$\min_{T_{in}} \left\{ h_n(\alpha_{1n}(T_{in}), \alpha_{2n}(T_{in}), \dots, \alpha_{In}(T_{in})) + g_n(\beta_{1n}(T_{in}), \beta_{2n}(T_{in}), \dots, \beta_{In}(T_{in})) \right\}, \quad (9)$$

where

$$h_n(x_1, x_2, \dots, x_I) = \sum_{i=1}^I x_i r_{in} \quad (10)$$

and

$$g_n(x_1, x_2, \dots, x_I) = J_{n+1}(x_1, x_2, \dots, x_I) + \sum_{i=1}^I (\rho_i + d_{in} - x_i) r_{in}. \quad (11)$$

If $\alpha_{in}(T_{in})$ and $\beta_{in}(T_{in})$ in (9) are replaced with the expressions obtained in equations (4)-(5), then the derivative of this function with respect to the upper acceptance limit U_{ink} is

$$- \frac{\partial h_n}{\partial x_i}(\alpha_{1n}, \alpha_{2n}, \dots, \alpha_{In}) p_{ink}(U_{ink}) f_{ink}(U_{ink}) \\ + \frac{\partial g_n}{\partial x_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In}) (1 - p_{ink}(U_{ink})) f_{ink}(U_{ink}),$$

where, to simplify notation, α_{in} stands for $\alpha_{in}(T_{in})$ and β_{in} stands for $\beta_{in}(T_{in})$. This expression equals zero if

$$p_{ink}(U_{ink}) = \frac{\frac{\partial g_n}{\partial x_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In})}{\frac{\partial h_n}{\partial x_i}(\alpha_{1n}, \alpha_{2n}, \dots, \alpha_{In}) + \frac{\partial g_n}{\partial x_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In})}. \quad (12)$$

The second derivative of the objective function (9) will be positive if $f'_{ink}(U_{ink}) < 0$ and $p'_{ink}(U_{ink}) < 0$. One would expect the second order conditions to hold, since the upper limit is at a point where both the frequency of measurement and the probability that a measurement corresponds to a valid board are decreasing.

The same argument can be used to show that the optimal lower acceptance limit satisfies

$$p_{ink}(L_{ink}) = \frac{\frac{\partial g_n}{\partial x_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In})}{\frac{\partial h_n}{\partial x_i}(\alpha_{1n}, \alpha_{2n}, \dots, \alpha_{In}) + \frac{\partial g_n}{\partial x_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In})}. \quad (13)$$

The second derivative of the objective function (9) will be positive if $f'_{ink}(L_{ink}) > 0$ and $p'_{ink}(L_{ink}) > 0$. Again, the second order conditions are consistent with our intuition.

From the definition of h_n and g_n in (10)-(11), it follows that

$$\frac{\partial h_n}{\partial x_i}(\alpha_{1n}, \alpha_{2n}, \dots, \alpha_{In}) = r_{in} \quad (14)$$

and

$$\frac{\partial g_n}{\partial x_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In}) = \frac{\partial J_{n+1}}{\partial \rho_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In}) - r_{in}. \quad (15)$$

Equations (12) and (13) now become

$$p_{ink}(L_{ink}) = p_{ink}(U_{ink}) = 1 - \frac{r_{in}}{\frac{\partial J_{n+1}}{\partial \rho_i}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In})}. \quad (16)$$

This expression has an intuitive meaning: the probability that a component is bad at the acceptance and rejection cutoff points should be equal to the marginal cost of a false rejection divided by the marginal cost of a false acceptance. As this cost ratio increases, it becomes less costly to accept parts, and the acceptance region is increased. Notice that the right side of (16) is a constant for each measurement $k = 1, \dots, K_{in}$, which we denote by C_{in} . The set of Pareto optimal testing policies can be generated by letting C_{in} vary from 0 to 1.

The existence of lower and upper limits L_{ink} and U_{ink} satisfying (16) is guaranteed by the continuity and unimodality assumptions of $f_{ink}(x)$ and $p_{ink}(x)$ stated in Section 2. We will not pursue weaker necessary and sufficient conditions for the existence of these optimal upper and lower limits, since in our case study these functions will be modeled by Gaussian density functions that satisfy the necessary conditions.

We can let the right side of (16) vary between 0 and 1, and use equations (1), (4), (5) and (16) to compute in advance all possible optimal testing policies, and generate the functions $\alpha_{in}(C_{in})$ and $\beta_{in}(C_{in})$. As a result, the minimization in the dynamic programming recursion (7) does not have to explicitly consider all possible testing policies, but only these two functions. The nonlinear curve in Figure 7 in Section 4.2 shows an example of the tradeoff curve of the expected number of type I errors per board ($\alpha_{in}(C_{in})$) versus the expected number of type II errors per board ($\beta_{in}(C_{in})$) obtained by letting C_{in} vary from 0 to 1. Since these quantities are independent of k , the curve was derived by aggregating all several hundred measurements (one per component) for this testing stage and defect type, which was the in circuit test that detects defective components.

We now turn to a special case where the set of Pareto optimal testing policies can be more explicitly described. This special case requires three further assumptions that are satisfied in our case study and often hold in practice. First, we assume that the measurement noise and the true value of the component being measured are normally distributed. It follows that $p_{ink}(x)$ can be simplified to (all subscripts ink are being dropped for better readability)

$$p(x) = \frac{1}{2} \left(\operatorname{erf} \left(\frac{(G_u - \mu_\xi)\sigma_e^2 + (G_u + \mu_e - x)\sigma_\xi^2}{\sigma_e\sigma_\xi\sqrt{2(\sigma_e^2 + \sigma_\xi^2)}} \right) - \operatorname{erf} \left(\frac{(G_l - \mu_\xi)\sigma_e^2 + (G_l + \mu_e - x)\sigma_\xi^2}{\sigma_e\sigma_\xi\sqrt{2(\sigma_e^2 + \sigma_\xi^2)}} \right) \right), \quad (17)$$

where:

μ_e and σ_e are the expected value and the standard deviation, respectively, of the measurement noise;

μ_ξ and σ_ξ are the expected value and the standard deviation, respectively, of the true values of the component being measured; we will also refer to μ_ξ as the *nominal value* of the component;

G_l and G_u are the lower and upper limits respectively of the interval G such that the measured component is good if its true value lies inside it;

erf is the error function, $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$, which is plotted in Figure 3.

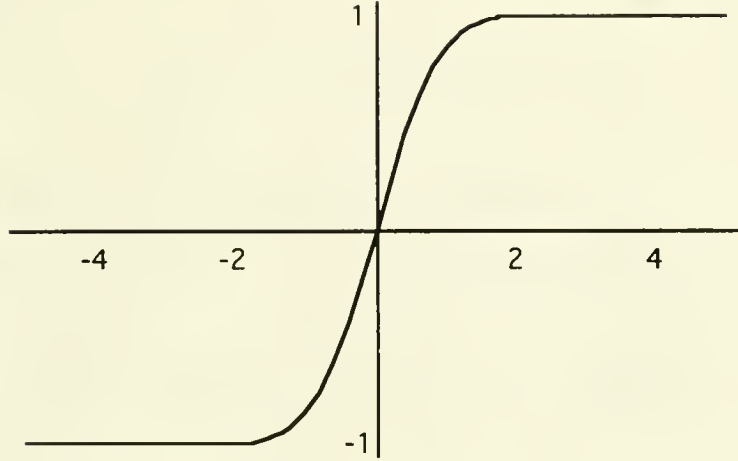


Figure 3. The error function.

We also assume that

$$\frac{|G_u - G_l|}{6\sigma_\xi} > 1 \quad (18)$$

and

$$\sigma_\xi > 3\sigma_\epsilon. \quad (19)$$

In quality management terminology, inequality (18) states that the *machine capability index* is greater than 1; that is, the natural process range, $6\sigma_\xi$, is smaller than the product specification range, $|G_u - G_l|$. Since many companies are currently striving for “6 σ ” capability (i.e., $|G_u - G_l| > 12\sigma_\xi$), (18) often holds in practice. Inequality (19) requires that the test measurements be reasonably reliable.

By conditions (18)-(19), one term of the right side of equation (17) will always be equal to 1 or -1, implying that $p(L) = p(U)$ for any L and U such that

$$L = \mu_\xi + \mu_\epsilon + (G_l - \mu_\xi)z \quad (20)$$

and

$$U = \mu_\xi + \mu_\epsilon + (G_u - \mu_\xi)z, \quad (21)$$

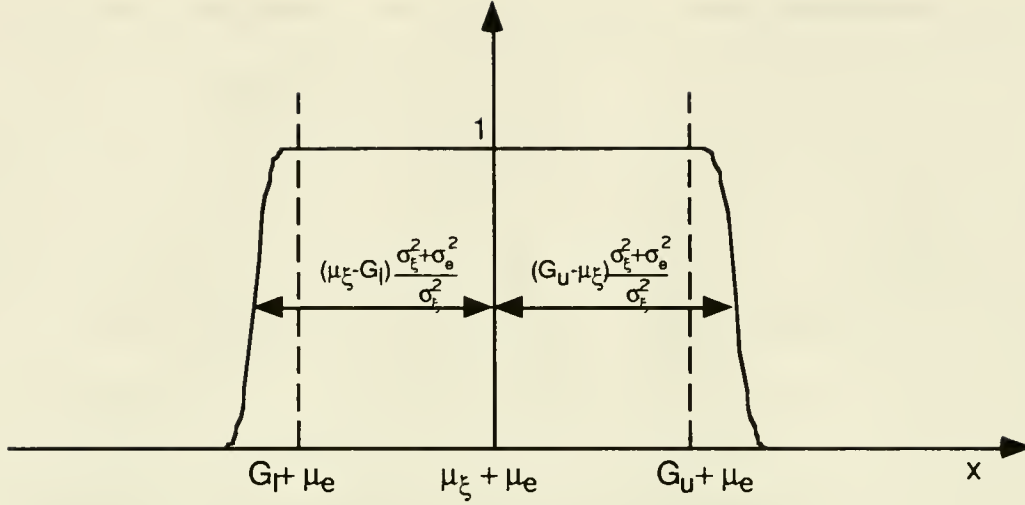


Figure 4. The function $p(x)$.

for any $z > 0$. Thus, the entire set of Pareto optimal policies are easily described and can be easily generated.

As a side comment, if the interval $[G_l, G_u]$ can be written as $[\mu_\xi - s, \mu_\xi + s]$, that is, the interval is centered around the nominal value, then the lower and upper limits to accept the component should be centered at the nominal value plus the expected value of the measurement noise. It can also be shown (see Chevalier 1992 for details) that this claim holds under the normality assumption, even if conditions (18)-(19) are not satisfied.

Figure 4 shows the form of the function $p(x)$ under assumptions (18)-(19). This function is equal to 1 between $G_l + \mu_e$ and $G_u + \mu_e$, and drops off very steeply to 0 somewhere outside this interval. The function is only symmetric if the tolerance interval is centered around the nominal value. Since the right side of (16) will take on values strictly between 0 and 1, the optimal cutoff points will most probably be located in the steep portions of the curve in Figure 4. Hence, the estimates $\hat{\mu}_\xi, \hat{\mu}_e, \hat{\sigma}_\xi, \hat{\sigma}_e$ used to obtain $\hat{p}(x)$, which is our estimate of $p(x)$, need to be very accurate. For example, suppose that we seek to find the two cutoff points such that $p(x) = 0.9$ and decide to use the cutoff points \bar{x} such that $\hat{p}(\bar{x}) = 0.9$. Furthermore, suppose that $\sigma_e = 0.9\hat{\sigma}_e$ (i.e., the estimate is 11.1% too high) and all other estimates are perfectly accurate. Under these assumptions, we used data from

various components in our study and found that $p(\bar{x})$ takes on values between 0.2 and 0.4. Unfortunately, despite gathering a large amount of data in the case study, we will see later that the 95% confidence intervals on the estimates used for $p(x)$ are much wider than this 11.1% overestimate. For this reason, it is extremely difficult to obtain an optimal testing policy that provides reliable performance.

Consequently, it does not seem worth the considerable effort required to derive the optimal testing policy. Instead, we propose to simply set the lower and upper cutoff limits to the middle points of the ascent and descent of $p(x)$ in Figure 4; that is, we find the two points such that $p(x) = 0.5$. These points, denoted by x_l and x_u , are derived by setting the argument of either error function in (17) equal to zero, which yields

$$x_l = \mu_\xi + \mu_\epsilon + (G_l - \mu_\xi) \left(\frac{\sigma_\xi^2 + \sigma_\epsilon^2}{\sigma_\xi^2} \right) \quad (22)$$

and

$$x_u = \mu_\xi + \mu_\epsilon + (G_u - \mu_\xi) \left(\frac{\sigma_\xi^2 + \sigma_\epsilon^2}{\sigma_\xi^2} \right). \quad (23)$$

Although the value in the right side of (16) will often be closer to 0.9 than 0.5 in practice, the corresponding difference in the cutoff points will typically be dwarfed by the inaccuracies in the measurement errors. Moreover, the cutoff points x_l and x_u have several desirable features. Even if the parameter estimates are inaccurate, the testing policy (x_l, x_u) is still Pareto optimal. Also, this policy is easy to find and can be expected to be close to the optimal policy if the estimates used for $p(x)$ are reasonably accurate. Furthermore, since the testing policy is derived in closed form, it is intuitively appealing, and is very easily understood and implemented by the test engineers. More specifically, expressions (22)-(23) have the following intuitive meaning (consult Figure 4): the acceptance interval should be shifted from the tolerance interval by the expected value of the measurement noise; the tolerance on both sides should be multiplied by a factor that is the ratio of the sum of variance of the true values of the component and the measurement noise over the variance of the true values of the component. This ratio shows very clearly how the acceptance interval

is affected by the measurement noise. For these reasons, x_l and x_u are used to define the testing policy in the case study described in the next section.

3.2. The Inspection Allocation Policy

The Pareto optimal curves characterized by $\alpha_{in}(C_{in})$ and $\beta_{in}(C_{in})$ can be substituted for $\alpha_{in}(T_{in})$ and $\beta_{in}(T_{in})$, respectively, into the dynamic programming optimality equation (7), and the optimal inspection allocation policy can be derived by solving the resulting dynamic program. However, the optimal solution becomes very tedious to calculate, since the functions $\alpha_{in}(C_{in})$ and $\beta_{in}(C_{in})$ and the I -dimensional state space are continuous, and I is of moderate size. More importantly, as pointed out in Section 3.1, the optimal testing policy is difficult to reliably implement in practice, and the testing policy derived in (22)-(23) has many desirable characteristics. Hence, we propose to use the testing policy defined in (22)-(23) and then find the optimal allocation policy given this fixed testing policy. Under this proposal, the dynamic programming equation (7) simplifies to

$$J_n(\rho_1, \rho_2, \dots, \rho_I) = \text{Min} \left[J_{n+1}(\rho_1 + d_{1n}, \rho_2 + d_{2n}, \dots, \rho_I + d_{In}), \right. \\ \left. t_n + \sum_{i=1}^I r_{in}(\rho_{i,n-1} + d_{in} + \alpha_{in} - \beta_{in}) + J_{n+1}(\beta_{1n}, \beta_{2n}, \dots, \beta_{In}) \right]. \quad (24)$$

where α_{in} and β_{in} are computed from equations (4)-(5) by setting the cutoff limits L and U to x_l and x_u .

Solving (8) and (24) numerically requires a discretization of the I -dimensional state space of incoming defects, and can be cumbersome in many cases. In contrast, an inspection allocation policy that solves (8) and (24) can be found relatively easily by employing an exhaustive search procedure over all 2^N policies, as described below.

Let π_n represent the inspection allocation policy at stage n , and let $\Pi_m = (\pi_1, \pi_2, \dots, \pi_m)$ denote the allocation policy for the line up to stage m . For $i = 1, \dots, I$ and $m = 1, \dots, N$, define

$$\delta_{i,\Pi_m} = \begin{cases} \beta_{im} & \text{if } \pi_m = \text{inspection,} \\ \delta_{i,\Pi_{m-1}} + d_{im} & \text{if } \pi_m = \text{no inspection,} \end{cases}$$

which is the expected number of defects of type i that leave stage m under policy Π_m . For $m = 1, \dots, N$, the cost c_{Π_m} of allocation policy Π_m is calculated from the cost $c_{\Pi_{m-1}}$ of allocation policy Π_{m-1} by

$$c_{\Pi_m} = \begin{cases} c_{\Pi_{m-1}} + t_m + \sum_{i=1}^I r_{im}(\delta_{i,\Pi_{m-1}} + d_{im} + \alpha_{im} - \beta_{im}) & \text{if } \pi_m = \text{inspection,} \\ c_{\Pi_{m-1}} & \text{if } \pi_m = \text{no inspection,} \end{cases}$$

where $c_{\Pi_0} = d_{i,\Pi_0} = 0$. The search procedure calculates the total inspection cost over the entire line, which is given by $c_{\Pi_N} + f \sum_{i=1}^I d_{i,\Pi_N}$, for all 2^N policies.

3.3. Dependencies Across Different Board Types

Until now, we have been considering only one type of circuit board. If different board types were completely independent, then all of our results would carry over. However, some tests actually measure several board types together, in which case either all the board types are tested or none are tested. Thus, the inspection allocation decisions taken independently might yield an infeasible solution. For shared tests that are nonparametric (i.e., have binary results), this dependence can be dealt with by employing a Lagrangian relaxation approach, where the total cost of the test is split among the different board types such that the solutions found independently coincide. To illustrate this approach, let us suppose that the total cost of a test that covers two board types is t . We run the model for the board types independently with the cost t split arbitrarily between the two types. If the decisions for the two types agree, then the solution is feasible and optimal. If the solutions do not agree, then modify the allocation of t by increasing the cost for the board type that is tested and decreasing the cost for the board type that is not tested. The problem is solved again for both board types, and if the solutions still do not agree, then we repeat the entire procedure until the solutions agree. Due to lack of data, this procedure could not be applied in our case study, and hence we are not in a position to report on the efficiency of any particular cost allocation algorithm. Notice that this approach becomes almost impossible to use with a parametric test, because the Lagrange multiplier must be a function of the testing policy that is used.

In the plant under study, the tests that measured more than one board type were the more comprehensive system tests, and hence were nonparametric.

Congestion also leads to dependencies across board types: if too many board types are tested at the same stage, then the testing load may be undesirably high at that stage. Moreover, if many of these board types are not tested at the previous stages, then many repairs will be required at that stage, which leads to additional work. At considerable computational expense, queueing effects could be incorporated here by a Lagrangian relaxation method similar to that described above, or by directly incorporating queueing costs (derived from a product form queueing network, for example) into the objective function. We did not pursue this avenue because the congestion effects at the Hewlett-Packard facility were negligible. Readers are referred to Tang (1991) for an inspection allocation problem with queueing costs.

3.4. *Improved Testing Equipment*

Our model explicitly takes into account the effect of measurement noise, which enables us to determine the value of improved testing equipment that possesses lower measurement noise. This issue is very important, since testing equipment is extremely expensive and it is not clear, a priori, how a reduction in measurement errors affects the overall cost of inspection.

To evaluate improved testing equipment in the context of our model, the new noise parameters μ_e and σ_e^2 are substituted into equations (22)-(23) and the exhaustive search procedure described in Section 3.2 is carried out. The optimal total inspection cost for the new equipment can then be compared with the corresponding value for the old equipment.

An experiment was performed to illustrate the magnitude of the savings achievable with more precise equipment. We assumed that a hypothetical new tester had a measurement noise whose standard deviation was half of that of the current tester. For the in circuit test that detects defective components in our case study, the tradeoff curve for the new and

original test equipment is pictured in Figure 5, which shows that both types of errors could simultaneously be divided by at least two using the new equipment; the upper curve in this figure is identical to the curve in Figure 7. If we assume similar error reductions for all components on the board, then the reduction in total inspection cost is on the order of 5 percent.

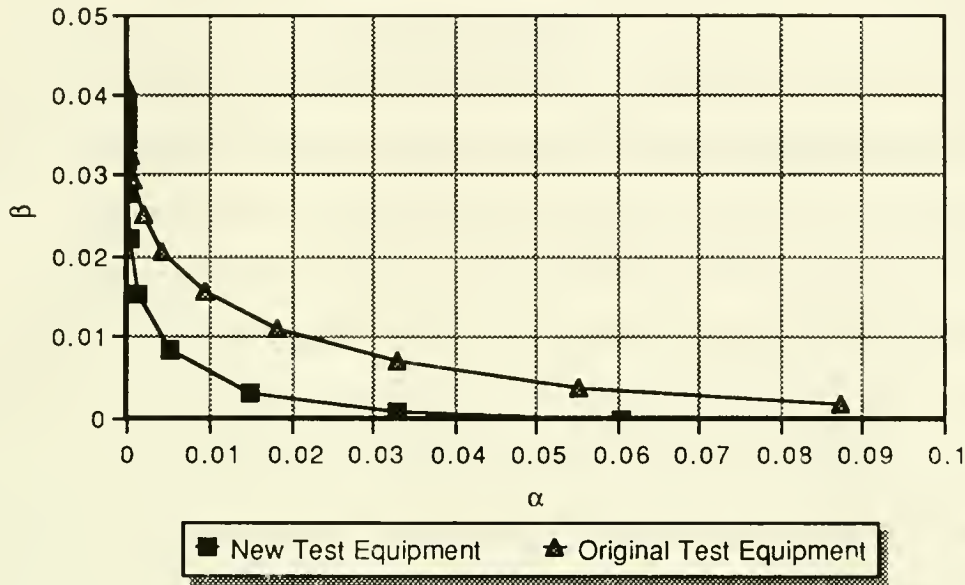


Figure 5. Reduction in testing errors.

3.5. Allocation of Quality Improvement Efforts

Engineering resources to work on quality improvements are limited, and consequently it is important to dedicate resources to projects that have the maximum impact. In a slightly different context, Albin and Friedman (1991) show that the choice of the most valuable project is not always straightforward. In particular, they show that the traditional Pareto chart is not an adequate tool when defects are clustered. It is also possible that a Pareto Chart would be misleading for the problem considered here, since different types of defects are detected in different proportions at different stages and hence the cost of different types of defects can be very different.

If the inspection policy for all stages following n is fixed, then the derivative of J_{n+1} with respect to its i^{th} argument is a constant for each i . This constant is the marginal expected cost of having one more defect of type i leave stage n and is easily calculated. Let q_{im} be the probability that a defect of type i arriving at stage m is detected at stage m . If we assume that all defects of type i reaching stage m are equally likely to be detected, then

$$q_{im} = \frac{\delta_{i,m-1} + d_{im} - \delta_{im}}{\delta_{i,m-1} + d_{im}}.$$

The probability that a defect of type i leaving stage n is repaired at stage $m > n$ is

$$\left(\prod_{l=n+1}^{m-1} q_{il} \right) (1 - q_{im}).$$

As a result, the expected total inspection cost created by this defect is

$$\frac{\partial J_{n+1}}{\partial \rho_i}(\cdot) = \sum_{m=n+1}^N \left[\left(\prod_{l=n+1}^{m-1} q_{il} \right) (1 - q_{im}) r_{im} \right] + \left(\prod_{l=n+1}^N q_{il} \right) f.$$

Hence, under any fixed inspection policy, this quantity is the marginal cost reduction obtained from reducing type i defects at stage n .

Recall that the hierarchical test coverage assumption implies that more defects become detectable at each inspection stage; this was modeled by assuming that, on average, d_{in} new type i defects appeared per board at stage n . As a result, working on the root cause of type i defects will simultaneously reduce the number of type i defects appearing at every stage. If we assume that the proportion of type i defects across the different stages does not change, then the marginal cost from reducing type i defects is

$$\frac{\sum_{n=1}^N \frac{\partial J_{n+1}}{\partial \rho_i}(\cdot) d_{in}}{\sum_{n=1}^N d_{in}}.$$

4. Case Study

We now describe the application of this model to the Hewlett-Packard facility. This plant has $N = 3$ inspection stages as shown in Figure 1. Approximately 50 different board

types are manufactured at this facility, and these boards go into various models of a single line of final products. The number of boards produced of each type is relatively low, ranging from 1,000 to 10,000 per year. Hence, tests must be flexible in order to accommodate many different board types. Figure 2 illustrates the basic structure of the tests performed. We grouped the many different categories of defects that are used internally into 7 broad types, based on their similarity for testing purposes. These seven categories are open and shorts, missing or wrong components, defective components, two different soldering defect types (one is much harder to detect than the other), miscellaneous, and testing problems. This last category consists entirely of false defects, and occurs when a defect is thought to be detected, but no problem can be found. Except for this last defect type, the number of measurements for each stage and defect type, K_{in} , ranges from about 20 to several hundred.

The nature of the final product, which cannot be revealed, requires very strict tolerances on the circuit boards and on their components, which partially explains why this facility currently tests all boards at all stages. The facility uses no consistent procedure for determining testing policies, and the policy is highly dependent on the particular engineer in charge of the test. The managers at this facility estimate that the total cost of inspection represents about half of the total manufacturing cost.

It would be an enormous task to collect data and derive the proposed inspection policy for all 50 board types. Instead, we applied our model in a very limited manner in order to exhibit its effectiveness, and left the appropriate software with the Hewlett-Packard engineers, who are carrying out the implementation. The data collection procedure is described in Section 4.1 and our results are given in Section 4.2.

4.1. *Data Collection*

Our model requires three different types of data. The first type is the cost parameters, and the second concerns the occurrence of defects, and the coverage and reliability of each inspection stage. The last and perhaps most difficult type of data to gather pertains to the

testing process.

Testing and repair are two activities that are closely linked, since they are performed by the same people in the same work area. Their cost includes technician time, floor space, equipment and overhead for management time. To allocate cost between these two activities, we use estimates of the time that different engineers and technicians spend on each of these activities. Both of these costs vary for each board type; in particular, they depend on the complexity and design of a board. These costs are proportional to the overall volume of inspection that takes place at a stage, and hence to the number of boards processed.

The procedure we used to estimate these cost parameters depended on the level of detail of the available data. For some stages, we obtained estimates of each of the different costs (engineering, equipment, floor space, etc.) individually. In these cases, we split each of the costs according to how much could be allocated either to testing (and testing maintenance) or repair (and diagnosis). The total cost for each of these activities was then divided by the total time the boards underwent testing and repair, respectively, to obtain a cost rate for each activity. Finally, multiplying the cost rate by the average time it took to test or repair a particular board type gave the corresponding cost estimate. In the cases where detailed overhead costs could not be obtained, we multiplied the average time to test and repair a particular board type by a flat overhead rate that included all the above costs. In the case study, the repair cost was assumed to be identical for each defect type for a particular board type. Although this is an approximation of reality, we did not have enough data to estimate these costs for each individual defect type. However, the test and repair costs did vary widely across board types and across inspection stages.

The cost of a defect on a board leaving the plant includes the cost of an on site repair, analysis and repair of the defective board at the plant, and the quality department's estimate of the cost of lost goodwill. We estimated the expected number of lost sales induced by a defective unit, and then set the goodwill cost equal to this quantity times the profit generated by one unit.

To estimate d_{in} , the expected number of new type i defects per board appearing at stage n , we used historical data about the total number of defects per board of each type detected at the different stages of inspection, ϕ_{in} . Then the engineers in charge of each test were asked to estimate the proportion a_{in} of defects of each type that they felt their test should detect, and the proportion b_{in} of defects of each type detected by the test that were actually good boards (false defects). Some of these estimates were based on experiments, whereas others were based on the experience of the engineers. We also used historical data to estimate η_i , which is the number of type i failures per board that occurred during the warranty period of the equipment.

From this data, we calculated the number of defects per board that become detectable at each stage n , which is

$$d_{in} = a_{in} \left(\sum_{m=1}^N \phi_{im}(1 - b_{im}) + \eta_i - \sum_{m=1}^{n-1} d_{im} \right) \text{ for } i = 1, \dots, I \text{ and } n = 1, \dots, N, \quad (25)$$

since a failure during the warranty period was considered to be caused by an undetected defect. For defect type $i = 1, \dots, I$ and stage $n = 1, \dots, N$, the historical expected number of type I errors per board is

$$\alpha_{in} = \phi_{in} b_{in}, \quad (26)$$

and the historical expected number of type II errors per board is

$$\beta_{in} = \sum_{m=1}^n (d_{im} - \phi_{im}(1 - b_{im})). \quad (27)$$

These historical estimates will be used later. Notice that the procedure culminating in (26)-(27) is much simpler and probably more reliable than attempting to estimate these historical quantities via (4)-(5).

To derive the testing policy, we need the input data G , $e(x)$ and $\xi(x)$ for each quantity measured at each stage. For a system to function properly, the true value of each quantity measured should be in the corresponding interval G_{ink} . It is often very hard to determine the exact limit values on a component that will ensure the proper functioning of the board.

This problem was avoided here by using the specifications to which the component was bought. Our justification is that even if a component is bought with tighter specifications than actually needed, a component that does not meet these specifications signals some kind of abnormality, such as physical damage or a poor soldering, that may cause a real defect after the equipment has been utilized for some time.

To estimate the distribution of the measurement noise $\epsilon(x)$ and the distribution of the true value of the quantity measured $\xi(x)$, we only have the distribution of the measured values at our disposal. This measured value, which is recorded by the testing device, is the sum of the true value of the component and the measurement noise. Unfortunately, neither of these two quantities can be estimated independently. The true values are almost impossible to measure, since most components used at this facility are surface mount components, which are extremely small and fragile. Estimating the measurement noise is also a delicate task since the distribution of this noise will depend on many things, such as the type of component that is being measured, how the measurement is *guarded* (guarding is the technique used to try to isolate a component from the rest of the circuit board) and the topology of the board. As a result, the measurement noise can only be determined via experiments for each different board type.

At our request, a controlled experiment was performed at the Hewlett-Packard facility to study the measurement noise at the in circuit test level. At this facility, several in circuit testers, also called *testheads*, are used in parallel, and boards are typically tested on the first available testhead. Consequently, we wanted to find out what portion of the measurement noise is attributable to variations across different testheads. The experiment repeated each of 79 measurements $K = 10$ times consecutively on $H = 3$ different testheads for $B = 10$ different boards of the same type, and generated nearly 24,000 data points. One measurement was taken from each of the 79 components on each board, which consisted of 28 resistors, 22 capacitors, 13 diodes, 12 transistors and 4 inductors. The measurement noise was modeled

by expressing the measurements y_{bhk} as

$$y_{bhk} = \mu + \tau_b + \theta_h + \psi_{bh} + \epsilon_{bhk} \quad \text{for } b = 1, \dots, B, \quad h = 1, \dots, H, \quad k = 1, \dots, K, \quad (28)$$

where:

μ is the reference value for the component being measured;

τ_b is the average deviation of the measurements taken on the b^{th} board, and has zero mean and standard deviation σ_τ ;

θ_h is the average deviation of measurements taken on the h^{th} head, and has zero mean and standard deviation σ_θ ;

ψ_{bh} is the average deviation of measurements taken on the h^{th} head and the b^{th} board, and has zero mean and standard deviation σ_ψ ;

ϵ_{bhk} is the residual variation of a measurement that cannot be explained by the testhead, the board, or the interaction between the testhead and the board, and has zero mean and standard deviation σ .

This model incorporates the implicit assumption that the residual noise ϵ has the same variance on all testheads. The data indicated that this assumption did not hold for all components, which suggests that the three testheads were not equally calibrated when the data was gathered. Without this assumption, the estimation of the model parameters would have been greatly complicated. Moreover, by only gathering data from three testheads, a good estimation of the distribution of the variance of the residual noise across different testheads was not possible.

We estimate μ by \bar{y} , which is the average of all measurements taken. The variance of ϵ is estimated by

$$\hat{\sigma}^2 = \frac{\sum_{b=1}^B \sum_{h=1}^H \sum_{k=1}^K (y_{bhk} - \bar{y}_{bh})^2}{BH(K-1)},$$

which is the variance of successive measurements of the very same component on the same testhead, and represents the precision limitation of the tester for the component. The vari-

ance of ψ is estimated by

$$\hat{\sigma}_{\psi}^2 = \frac{\sum_{b=1}^B \sum_{h=1}^H (\bar{y}_{bh} - \bar{y}_b - \bar{y}_h - \bar{y})}{(B-1)(H-1)} - \frac{\sigma^2}{K},$$

and the variance of θ is estimated by

$$\hat{\sigma}_{\theta}^2 = \frac{\sum_{h=1}^H (\bar{y}_h - \bar{y})^2}{H-1} - \frac{\sigma_{\psi}^2}{B} - \frac{\sigma^2}{BK}.$$

These two quantities represent the amount of variation that is linked to the variations between testheads. One might think that a fixed effects model would be more appropriate for the testheads, since the facility uses only a fixed number of different testheads. But the variation between testheads evolves over time as a result of usage, maintenance, calibration, etc., and thus the random effects model seems more appropriate.

The average of all measurements taken on the b^{th} board equals $\mu + \tau_b$, and τ_b can be considered to be the variation in the measured values associated with the individual components on the b^{th} board. The variance of τ is estimated by

$$\hat{\sigma}_{\tau}^2 = \frac{\sum_{b=1}^B (\bar{y}_b - \bar{y})^2}{B-1} - \frac{\sigma_{\psi}^2}{H} - \frac{\sigma^2}{HK}.$$

Finally, the estimated parameters from the statistical model in equation (28) are used to estimate the parameters of the distributions of the measurement noise and the component values. We assume that μ_{ξ} , the nominal value of the component, is known (e.g., a 100 ohm resistor has a nominal value of 100 ohms), and let

$$\hat{\mu}_{\epsilon} = \bar{y} - \mu_{\xi}, \quad (29)$$

$$\hat{\sigma}_{\epsilon}^2 = \hat{\sigma}^2 + \hat{\sigma}_{\psi}^2 + \hat{\sigma}_{\theta}^2 \quad (30)$$

and

$$\hat{\sigma}_{\xi}^2 = \hat{\sigma}_{\tau}^2. \quad (31)$$

For each of the five component types, the Appendix contains tables of the estimates that were derived from this experiment. The results from this experiment were used to test

assumptions (18)-(19), which were required to derive the proposed testing policy, x_l and x_u . By substituting $\hat{\sigma}_\xi^2$ for σ_ξ^2 in (18), we found that the machine capability index was greater than one for all 79 components. To assess the validity of (19), readers are referred to the Appendix. For all four inductors in this sample, the estimated variance of the measurement noise, $\hat{\sigma}_e^2$, is higher than the estimated variance of the true values of these components, $\hat{\sigma}_\xi^2$. Hence, these measurements cannot distinguish between good and defective components. The estimated variance of the measurement noise is the same order of magnitude as the estimated variance of the true values for the 13 diodes, and none of the inductors or diodes satisfy (19). In contrast, for the other three component types, 50 of the 62 components, and hence 50 of 79 in total, satisfy (19). For these other three component types, most of the components have an estimated noise variance much smaller than the estimated variance of their true values. These components can be tested with good accuracy. This analysis has important practical implications: some components should not be tested (or, alternatively, a new test should be devised for them).

An important question for the design of future experiments is whether the variance of the noise associated with the different testheads can be predicted. It is much easier to run experiments on a single testhead than on several testheads, because in the latter case, experiments are much more difficult to schedule into the production process. To address this question, a regression was run to predict σ_e from σ and the absolute value of μ_e . The coefficients of correlation were consistently high; for example, $r^2 = 0.92$ for the resistors and $r^2 = 0.98$ for the capacitors.

We also gathered data from another board type during an actual production run, and used this data for two purposes. The data was first used to investigate the normality assumptions on the measurement error and true measurement value, which were also required to derive x_l and x_u . These assumptions are difficult to validate because the two distributions are not at our disposal. However, we did assess the normality of the measurement distribution, which is the convolution of the two distributions in question. Notice that the

results from the controlled experiment are not appropriate for testing this assumption, since the experiment contains many repeated measurements. The production data consists of a measured value from each of 350 components on 80 different boards. For each of the 350 components, we applied the goodness-of-fit test associated with the kurtosis measurement in equation (27.6a) of Duncan (1986) to the 80 measurement values. For more than 250 of the components, the normality assumption was not rejected at the 95% significance level.

The controlled experiments compete with the regular production runs for scarce resources, and the facility's emphasis on product output makes it impractical to run a similar controlled experiment for each of the 50 board types. Hence, the production data was also used to assess the qualitative similarity between the two board types. More specifically, the production data enabled us to estimate the total variance of the measurements, $\sigma_\epsilon^2 + \sigma_\xi^2$, and Appendix B of Chevalier (1992) displays the distribution of the total standard deviation for the different components of each type on the board. We found that these results and the corresponding results in our Appendix are strikingly similar. For example, most resistors have a total standard deviation near 0.3% in both appendices, and most capacitors have a total standard deviation between 2 and 4% in both cases. These results are very encouraging, particularly since the two board types plug into different subsystems of the final product, but a more thorough investigation of this issue is still needed. The controlled experiment described earlier is required to estimate the parameters σ_ϵ^2 and σ_ξ^2 individually, and this experiment should be repeated on some other board types before concluding that the parameters do not vary significantly by board type.

Unfortunately, a much larger experiment than the one performed here is required to obtain reliable estimates for σ_ϵ^2 and σ_ξ^2 ; the expected value of the measurement noise was easier to estimate. The 95% confidence interval for a typical component was such that the estimate could be off by about 50% in either direction. Comparing this with the example of a 11.1% overestimation described in Section 3.1, we see that these estimates are not reliable enough to use with any degree of confidence. To get out of this quandary, we attempted

to group components in subsets that had very similar properties. The key characteristic we used for aggregating was that the standard deviation of the measurement noise and the standard deviation of the component values represented nearly the same percentage of the nominal value of the component. Not only will the grouping of operations of different components increase the precision of the estimates for σ_e^2 and σ_ξ^2 , but it will also make the tradeoff curve in Figures 5 and 7 less tedious to calculate. Although our aggregation led to tighter confidence intervals for σ_e^2 and σ_ξ^2 , the intervals were still too large to implement the optimal testing policy in a reliable manner.

4.2. Results

As mentioned above, we did not apply the model to the entire Hewlett-Packard facility. The model was first used to find the optimal inspection allocation policy given the facility's *historical* testing policy. For this purpose, three types of circuit boards were chosen that were representative of the variety of boards manufactured; these board types are different than the types analyzed in the Appendix or Chevalier's Appendix B. Since the analog or digital nature of a board is one of its distinguishing features, we chose one board type with mostly digital components, one with mostly analog components, and one with a mixture of components. These board types were produced in volumes that were typical for the facility, and their yield ranged from relatively low to relatively high. Also, these three board types did not share any test and thus can be considered independently.

The optimal testing policy was derived for the in circuit test that detects defective components, using the board type analyzed in the Appendix. Hence, 79 (x_l, x_u) pairs were calculated, and the resulting expected number of type I and type II errors were compared to the corresponding quantities under the facility's current testing policy.

Optimal allocation policy. Figure 6 displays the frequency of the different defect types detected at each stage for the three board types under consideration. Each defect type is represented by the same pattern on all three charts. The values on each chart are arbitrary

in order to disguise the data, but the relative values across the three charts are approximately correct; these values are proportional to the quantities ϕ_{in} defined in Section 4.1. The figure shows that the number of defects and the predominant types of defects vary significantly across the different board types. For example, type 3 boards have roughly three times as many defects as type 1 board types, and the medium gray defect type is predominant on type 1 boards but hardly present on type 2 boards.

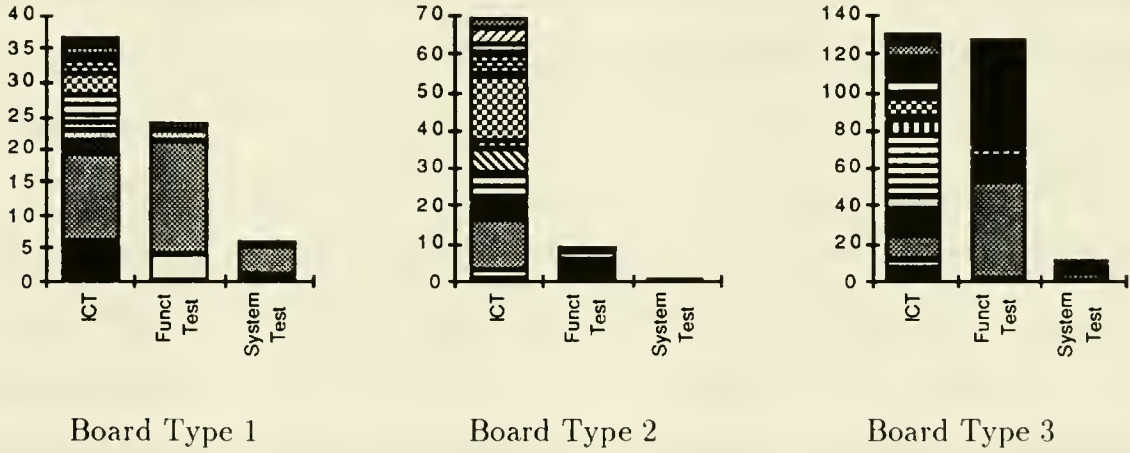


Figure 6. Frequency of defects detected at each stage for each board type.

The costs and the quantities d_{in} , α_{in} and β_{in} in (25)-(27) were estimated for each board type, and also varied considerably across board types. Recall that the quantities in (26)-(27) represent the expected number of errors per board under the facility's historical testing policy. These estimates were used in the exhaustive search procedure described in Section 3.2 to calculate the optimal inspection allocation for each board type, which is displayed in Table I. As expected, the optimal allocation varies for the different board types. Referring to Figure 6 and Table I, we see that the system test is only bypassed by board type 2, which has the lowest frequency of defects detected at system test. Since the total inspection cost represents about half of the total manufacturing cost, the savings realized by the optimal policy in the different cases are significant.

Testing Policy. The in circuit test for the board type considered in the Appendix tests for the six different defect types listed earlier. We focus on the test for defective components.

Table I. Optimal policies and savings.

Board Type	Current Inspection Allocation	Optimal Inspection Allocation	Cost Reduction
1	1-2-3	1-3	20%
2	1-2-3	2	23%
3	1-2-3	2-3	6.5%

where a measurement is taken from each of $K_{in} = 79$ components. The data in the Appendix was used to derive the estimates in (29)-(31) for each component. These quantities, along with μ_ξ and the interval $[G_l, G_u]$, allow us to determine via (22)-(23) the proposed testing policy (x_l, x_u) for each component. For example, the interval $[G_l, G_u]$ for component R168 in Table II is $[99, 101]$, and hence $x_l = 98.8847$ and $x_u = 100.9187$ for this component. Under normality assumptions, the expected number of type I and type II errors per board were calculated using (4)-(5). The resulting pair of values, $\alpha_{in}(T_{in})$ and $\beta_{in}(T_{in})$, correspond to the point a in Figure 7. The remainder of the Pareto optimal tradeoff curve in Figure 7 was derived by performing a similar aggregation using (20)-(21).

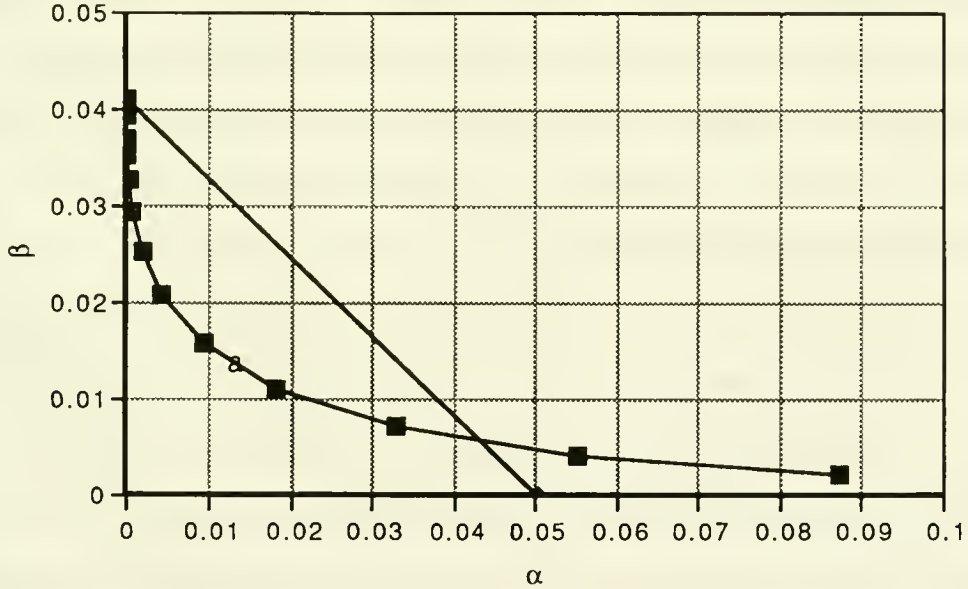


Figure 7. The current testing policy and the Pareto optimal tradeoff curve.

Hewlett-Packard engineers have observed that about 55% of the rejected components at the in circuit test are good components. Further analysis found that this percentage did not vary significantly by board type. To illustrate how this performance compares to that of our proposed testing policy, Figure 7 also shows a straight line that represents the set of policies such that 55% of the rejected components are good components. If we had obtained somewhat different estimates for σ_e^2 and σ_ξ^2 , then the proposed policy would have been on the tradeoff curve in the neighborhood of a , and a significant improvement over the current policy would still be achieved. The percentage improvement in Figures 5 and 7 appear to be of the same order of magnitude, which represents about a 5% reduction in inspection costs if implemented systemwide.

5. Concluding Remarks

The model presented here is built on two key features: hierarchical test coverage and measurement errors in the testing process. Although this model was developed in the framework of circuit board assembly, these features are present in many other settings. Hierarchical test coverage is a very natural property when inspection is performed at different stages in a manufacturing process; the increasing test coverage is then a direct consequence of the fact that potential defects are added between inspection stages by the intermediate manufacturing operations. Also, test measurements are typically subject to some noise, particularly in the manufacturing of high precision products.

For the problem considered here, an optimal inspection policy could be pursued for a modest size problem by discretizing the I -dimensional state space of defect types and the curves that can be generated by the Pareto optimal policies in (20)-(21). However, the case study reveals that the required precision in the parameter estimates are very difficult to achieve, even after performing controlled experiments and aggregating similar component types. As a result, we focus on developing a simple but systematic procedure for deriving the cutoff limits that characterize the testing policy.

For three representative board types in the case study, the optimal inspection allocation policy achieves a 10% to 20% reduction in expected inspection costs relative to the facility’s historical inspection policy, under the facility’s historical testing policy. For the in circuit test that detects defective components, the proposed testing policy significantly outperforms the facility’s historical testing policy on one board type, representing roughly a 5% reduction in inspection costs. Since the cost of inspection represents about half of the total direct manufacturing cost, these cost savings are significant. Moreover, both policies are relatively easy to implement in practice.

The Hewlett-Packard facility is just beginning to apply our model. Since a thorough investigation of measurement errors had never been undertaken at this facility, perhaps the biggest contribution thus far has been the statistical analysis of the data. We encountered many cases, previously unbeknownst to the facility’s engineers, where the mean measurement error was an order of magnitude larger than the variance of the measurement error; see the tables in the Appendix for some examples. These tables are extremely useful for developing some quick insights into the optimal testing policy. For example, resistors such as R158, R170 and R317 should perhaps not be tested (or at least have extremely slack cutoff limits), since the measurement noise is so large relative to the true component noise. In contrast, fairly tight cutoff limits can be set for most of the resistors, where $\sigma_\epsilon/\sigma_\xi$ is very small. The facility has begun implementing the proposed testing policy, but has not yet implemented the optimal inspection allocation policy. Although they plan to implement the latter policy, the different inspection stages are managed by different people, and organizational barriers still need to be overcome.

We started from a real problem at a specific industrial facility, and developed and analyzed a mathematical model to address this problem; however, the model is still relatively crude. Nevertheless, it is hoped that this work will help future researchers find better models for the large number of problems of this nature that are beginning to emerge. With the higher performance that is sought from manufactured products, it seems very likely that

inspection will become an increasingly important aspect of production management. With respect to model refinement, we believe that a key shortcoming in this model is the assumed independence between the distribution of measurement values at a given inspection stage and the inspection policy used at previous stages. A natural first step towards modeling this dependence is to assume that the true measurement value v_{ink}^t depends on the inspection policy at previous stages only through $\delta_{i,n-1}$, which is the expected number of type i defects present on a board at stage n . Under this assumption, the dynamic programming framework and the derivation of the Pareto optimal testing policies still hold; the only difference is that the quantities α_{in} and β_{in} in (7) and (16) are now functions of $\delta_{i,n-1}$, as well as the testing policy T_{in} . All that remains is the formidable task of estimating the influence of $\delta_{i,n-1}$ on v_{ink}^t from available data.

Acknowledgment

We are grateful to the people at the Hewlett-Packard plant in Andover, MA for their advice and their tremendous effort in helping us obtain the necessary data for our model. We also thank Arnie Barnett and Steve Pollock for helpful discussions. This research is supported by a grant from the Leaders for Manufacturing program at MIT and National Science Foundation Grant Award No. DDM-9057297.

References

- Albin, S. L. and D. J. Friedman. 1991. Off-Line Quality Control in Electronics Assembly: Selecting the Critical Problem. Working Paper, Rutgers University, Piscataway, NJ.
- Chevalier, P. B. 1992. Two Topics in Multistage Manufacturing Systems. Ph.D. dissertation, Operations Research Center, MIT, Cambridge, MA.

- Duncan, A. J. 1986. *Quality Control and Industrial Statistics*. Irwin, Homewood, IL.
- Eppen, G. D. and E. G. Hurst. 1974. Optimal Allocation of Inspection Stations in a Multistage Production Process. *Management Science* **20**, 1194-1200.
- Garcia-Diaz, A., J. W. Foster, and M. Bonyuet. 1984. Dynamic Programming Analysis of Special Multistage Inspection Systems. *IIE Transactions* **16**, 115-125.
- Lindsay, G. F. and A. B. Bishop. 1967. Allocation of Screening Inspection Effort—A Dynamic-Programming Approach. *Management Science* **10**, 342-352.
- Raz, T. 1986. A Survey of Models for Allocating Inspection Effort in Multistage Production Systems. *Journal of Quality Technology* **18**, 239-247.
- Tang, C. S. 1991. Designing an Optimal Production System with Inspection. *European Journal of Operations Research* **52**, 45-54.
- Villalobos, J. R., J. W. Foster, and R. L. Disney. 1992. Flexible Inspection Systems for Serial Multi-Stage Production Systems. Working Paper. Department of Industrial Engineering, Texas A&M University, College Station, TX.
- White, L. S. 1966. The Analysis of a Simple Class of Multistage Inspection Plans. *Management Science* **9**, 685-693.
- Yum, B. J. and E. D. McDowell, E. D. 1981. The Optimal Allocation of Inspection Effort in a Class of Nonserial Production Systems. *IIE Transactions* **13**, 285-293.

Table II. Resistors.

Component	Nominal Value (Ω)	Mean Measurement Error μ_e (%)	Measurement Error Std. Dev. σ_e (%)	Component Value Std. Dev. σ_ξ (%)
R158	10	1.6580	0.7778	0.3208
R160	21.5	0.7194	0.0224	0.2938
R102	31.6	0.3621	0.0655	0.2384
R168	100	-0.0983	0.0370	0.2835
R105	215	-0.1170	0.0157	0.1758
R106	1000	0.0142	0.0118	0.2279
R111	1000	0.0530	0.0421	0.2449
R113	1000	0.0965	0.0177	0.2088
R120	1000	0.1492	0.0118	0.1502
R162	1000	0.1059	0.0120	0.1192
R210	1000	0.0582	0.0116	0.1718
R211	1000	0.0280	0.0112	0.3055
R213	1000	0.0896	0.0125	0.2082
R214	1000	0.0408	0.0126	0.2850
R215	1000	-0.0084	0.0130	0.2286
R132	1780	1.9721	0.0700	0.2187
R117	2150	0.0126	0.0123	0.1623
R301	4640	-0.0987	0.0101	0.1435
R322	5110	-0.4706	0.0792	0.2541
R310	8250	-0.1981	0.0400	0.1090
R303	10000	-0.1581	0.0111	0.1422
R316	14700	-0.1642	0.0313	0.3167
R170	17800	-2.2224	0.2614	0.4085
R302	21500	-5.1729	0.2507	1.6664
R110	31600	0.0407	0.0162	0.1472
R321	68100	-0.2850	0.0608	0.2262
R133	121000	-0.1053	0.0269	0.2252
R317	464000	0.0122	0.1526	0.2318

Table III. Inductors.

Component	Nominal Value (μH)	Mean Measurement Error μ_e (%)	Measurement Error Std. Dev. σ_e (%)	Component Value Std. Dev. σ_ξ (%)
L101	1	37.5506	6.3128	1.1392
L102	1	46.9465	3.7830	1.8730
L103	1	45.8132	3.6897	1.7435
L104	1	37.0869	9.6679	0.2264

Table IV. Capacitors.

Component	Nominal Value (μF)	Mean Measurement Error μ_e (%)	Measurement Error Std. Dev. σ_e (%)	Component Value Std. Dev. σ_ξ (%)
C128	0.01	-2.2309	0.9593	1.3864
C125	0.1	2.9324	0.3602	1.8012
C201	0.1	1.3018	0.3156	5.2466
C202	0.1	1.0444	0.3189	4.6113
C203	0.1	2.5904	0.2824	3.5374
C204	0.1	-0.3741	0.3247	4.2337
C205	0.1	2.5666	0.3238	3.0688
C308	0.1	2.8296	0.4630	4.2328
C305	0.3	1.9978	5.9763	2.8549
C301	0.4	-2.8866	1.2112	2.3799
C114	0.47	-12.5308	1.4739	6.3439
C126	0.47	0.8126	0.4203	2.2399
C129	0.47	-0.2691	0.2845	1.7317
C130	0.47	0.7670	0.3225	1.4685
C309	0.47	-0.3155	0.1996	1.8899
C310	1	-7.6399	0.1894	0.5920
C127	6.8	-1.5503	0.1315	1.6210
C107	6.9	2.3444	0.2334	1.9289
C118	6.9	-2.1074	0.1436	1.1643
C120	7.2	-0.1760	0.1595	1.4322
C101	33	0.1986	0.5414	1.9831
C131	33	-0.0072	0.2814	2.6723

Table V. Transistors.

Component	Nominal Value (V)	Mean Measurement Error μ_e (%)	Measurement Error Std. Dev. σ_e (%)	Component Value Std. Dev. σ_ξ (%)
Q104	50.1388	N/A	0.3870	0.7738
Q105	59.753133	N/A	9.6449	4.3180
Q108	92.116	N/A	0.4122	2.2315
Q109	37.7534	N/A	0.3575	0.9633
Q110	28.299333	N/A	0.5867	5.4653
Q113	28.703933	N/A	0.5577	2.2126
Q114	36.0998	N/A	0.2703	8.3509
Q118	36.404233	N/A	0.3368	5.5723
Q119	37.419267	N/A	0.2479	1.6809
Q120	34.036667	N/A	0.6842	15.5969
Q122	48.807333	N/A	0.3554	1.9639
Q124	75.055467	N/A	1.5632	0.4967

Table VI. Diodes.

Component	Nominal Value (V)	Mean Measurement Error μ_e (%)	Measurement Error Std. Dev. σ_e (%)	Component Value Std. Dev. σ_ξ (%)
CR103	1.622672	N/A	0.0890	0.1148
CR104	0.70446	N/A	0.1272	0.1220
CR105	0.724097	N/A	0.1102	0.0556
CR106	0.726004	N/A	0.1210	0.1145
CR107	0.72484	N/A	0.1269	0.0821
CR108	0.723124	N/A	0.1144	0.0904
CR109	0.723348	N/A	0.1136	0.0695
CR111	2.168599	N/A	0.1133	0.2104
CR301	0.717516	N/A	0.1131	0.2404
CR302	0.590526	N/A	0.1536	0.2206
CR303	0.59102	N/A	0.1833	0.1520
CR304	0.595119	N/A	0.1643	0.1513
CR305	0.597192	N/A	0.1566	0.1970

Date Due

JUN 31 1894

Lib-26-67

MIT LIBRARIES



3 9080 00719471 2

